ISSN: 2320-5407

International Journal of Advanced Research

Publisher's Name: Jana Publication and Research LLP

www.journalijar.com

REVIEWER'S REPORT

Manuscript No.: IJAR-52079 Date: 05-06-2025

Title: Hyperparameter Impact on Learning Efficiency in Q-Learning and DQN Using OpenAI Gymnasium Environments

Recommendation:
Accept as it is.................**YES............**
Accept after minor revision.................
Accept after major revision ....
Do not accept (*Reasons below*) ....

Rating	Excel.	Good	Fair	Poor
Originality			√	
Techn. Quality			√	
Clarity		√		
Significance			√	

Reviewer's Name: Mr Bilal Mir

Reviewer's Decision about Paper: **Recommended for Publication.**

Comments (*Use additional pages, if required*)
.

Reviewer's Comment / Report

1. Relevance and Contribution

This study presents a highly relevant exploration within the domain of reinforcement learning (RL), specifically examining how hyperparameter configurations influence the learning performance of two foundational RL algorithms—Q-Learning and Deep Q-Network (DQN). The research is contextualized using the CartPole-v1 environment from OpenAI Gymnasium, a standard and widely accepted testbed for benchmarking RL algorithms. The focus on hyperparameter tuning reflects an important and ongoing concern in practical RL applications, enhancing the paper's contribution to both theoretical understanding and empirical practice.

2. Abstract

The abstract succinctly summarizes the comparative analysis, experimental methodology, and

ISSN: 2320-5407
International Journal of Advanced Research
Publisher's Name: Jana Publication and Research LLP
www.journalijar.com

REVIEWER'S REPORT

principal outcomes. It clearly communicates that DQN significantly outperforms Q-Learning in terms of training rewards and test performance. The abstract also highlights specific hyperparameter values (e.g., learning rate and batch size) contributing to DQN's success, while noting Q-Learning's limitations due to discretization and hyperparameter sensitivity. The content is concise and informative, providing a strong overview for readers.

3. Introduction

The introduction is thorough, effectively framing the significance of RL and the roles of Q-Learning and DQN. It successfully articulates the theoretical distinctions between the two algorithms, particularly in how they address state representation—tabular for Q-Learning versus continuous for DQN via neural networks. The discussion on the importance of hyperparameters, including learning rate, discount factor, and exploration strategies, adds strong contextual grounding for the study's objectives.

The choice of CartPole-v1 as the benchmark environment is well-justified, and prior findings are logically integrated to underscore the gap this research aims to fill. The introduction strikes a balance between conceptual depth and accessibility, making it suitable for both novice and advanced RL researchers.

4. Methodological Approach

While detailed methods are not included in the provided excerpt, the described approach suggests a controlled and empirical analysis of key hyperparameters across both algorithms. The mention of visualizations, training and testing reward summaries, and the use of performance metrics (such as peak reward and convergence rate) suggests a structured experimental design. The study's focus on both convergence behavior and final reward performance ensures a comprehensive understanding of algorithmic efficiency.

5. Analytical Rigor and Findings

The results highlighted in the abstract and introduction present a strong narrative. The stark contrast in performance—DQN reaching the maximum reward of 500 versus Q-Learning's average reward of 23—clearly demonstrates the differing capabilities of the algorithms under variable hyperparameter conditions. Furthermore, the identification of learning rate and batch size as key factors in DQN's success is valuable. The assertion that inappropriate

ISSN: 2320-5407
International Journal of Advanced Research
Publisher's Name: Jana Publication and Research LLP
www.journalijar.com

REVIEWER'S REPORT

hyperparameters can lead to stagnation or policy inefficiency reinforces the practical necessity of tuning in RL implementations.

6. Thematic Coherence and Depth

The paper successfully connects theoretical discussion with empirical results. It engages with the broader RL community by addressing a persistent challenge: translating algorithmic potential into practical performance through hyperparameter optimization. The relationship between reward outcomes and parameter tuning is well-articulated, adding to the coherence and relevance of the study.

7. Overall Evaluation

This is a focused, well-structured, and academically relevant study. It presents a clear comparison of Q-Learning and DQN in a standard environment, effectively highlighting the significant influence of hyperparameter choices on learning efficiency. The clarity of purpose, empirical grounding, and direct applicability to reinforcement learning practices make it a valuable contribution to the field.

Verdict:

A technically sound and well-motivated study that offers insightful empirical evidence on the role of hyperparameters in shaping reinforcement learning outcomes.